基于多权值的 Slope One 协同过滤算法*

覃幸新　王荣波　黄孝喜　谌志群

(杭州电子科技大学认知与智能计算研究所　杭州　310018)

摘要: 【目的】针对 Slope One 算法未考虑项目相似性、项目属性和对目标用户已有评分同等考虑进而导致推荐准确度降低的问题进行改进。【方法】提出一种基于改进的项目相似性度量、改进的项目属性相似性度量和用户评分概率函数的多权值的Slope One协同过滤算法, 在项目相似性度量方面将共同评价的两个项目的用户数量和Pearson 相关系数相融合, 在项目属性相似性度量方面将修正的拉普拉斯平滑与 Jaccard 系数相结合, 同时利用用户评分概率函数对用户已有评分进行有效区分。【结果】实验结果表明, 本文方法相比于原 Slope One 算法, MAE 值下降了 5.4%, 能够获得更好的推荐准确度。【局限】只关注推荐系统中用户对项目产生的评分, 并没有关注用户对项目给出的评论, 在一定程度上影响了推荐效果。【结论】本文方法更能适应评分数据稀疏性, 有效提高了推荐系统的推荐质量。

关键词: 协同过滤　Slope One　多权值　项目相似性　项目属性

分类号: TP391　G35

1　引　言

随着现代信息技术的迅猛发展, 互联网信息充斥着人们的生活, 人们从信息匮乏时代走向了信息过载时代。从海量的用户历史数据中提取出有价值的信息变成如今信息时代的难题, 推荐系统正是旨在解决这一难题而逐渐成为一个热门的研究课题。协同过滤(Collaborative Filtering, CF)算法是推荐系统中主流的算法之一[1-2], 其中基于内存的协同过滤算法分为基于用户的协同过滤算法(User-based Collaborative Filtering, UCF)和基于项目的协同过滤算法(Item-based Collaborative Filtering, ICF)两种[3]。基于项目的协同过滤算法则是通过对项目进行相似性度量进而寻找最近邻, 基于项目最近邻将类似的物品推荐给用户。本文研究的 Slope One 算法(Slope One Algorithm, SO)是 Lemire 等[4]提出的一种基于项目的协同过滤算法, 但是和传统基于项目的协同过滤算法不同的是, Slope One 算法通过计算项目之间的平均评分偏差并结合目标用户已有评分进行评分预测。相比上述两种传统的协同过滤算法, 具有运用简单快捷、对新用户和新项目适应性强和运行效率高等优点。

本文针对 Slope One 算法未考虑项目相似性、项目属性和对目标用户已有评分同等考虑进而导致推荐准确度降低的问题, 提出一种基于改进的项目相似性度量、改进的项目属性相似性度量和用户评分概率函数的多权值的 Slope One 协同过滤算法, 使得具有不同项目相似度的项目数据、不同的项目属性相似度和不同的目标用户评分概率以不同的权值参与到评分预测中, 提高了推荐准确度。

2　相关工作

2.1　Slope One 算法

Slope One 算法是一种简洁且高效的评分预测算法, 其核心思想是采用线性回归方法进行评分预测, 其线性回归公式表示为 $f(x)=x+b$[4], 其中, x 为在推荐系统中目标用户对项目产生过的历史评分, b 为项目之

通讯作者: 覃幸新, ORCID: 0000-0002-9378-2333, E-mail: 15957191408@163.com。

*本文系国家自然科学基金青年基金项目"基于马尔科夫树与 DRT 的汉语句群自动划分算法研究"(项目编号: 61202281)和教育部人文社科规划青年基金"基于语义相关性的汉语组块切分模型研究"(项目编号: 12YJCZH201)的研究成果之一。

间的平均评分偏差。对于评分矩阵中不同的项目 i 和 j，它们的平均评分偏差 dev_{ij} 计算方法如公式(1)所示。

$$dev_{ij}=\frac{\sum_{u\in U_{ij}}(r_{ui}-r_{uj})}{|U_{ij}|} \tag{1}$$

其中，r_{ui} 表示用户 u 对项目 i 的评分，U_{ij} 表示对项目 i 与 j 都产生了评分的用户集合，$|U_{ij}|$表示集合 U_{ij} 中的用户个数。

利用公式(1)求得两个项目的平均评分偏差 dev_{ij} 后，用 pre_{ui} 表示使用 Slope One 算法进行评分预测得到的用户 u 对项目 i 的评分，其计算方法如公式(2)所示。

$$pre_{ui}=\frac{\sum_{j\in I(u)}(r_{uj}+dev_{ij})}{|I(u)|} \tag{2}$$

其中，$I(u)$表示用户 u 产生过评分，且满足($i\neq j$, $|U_{ij}|>0$)的项目集合。

2.2 Weighted Slope One 算法

Slope One 算法在计算项目 i 相对于项目 j 的平均评分偏差 dev_{ij} 时没有考虑到使用对两个项目都产生过评分的用户数量来调整 dev_{ij} 在评分预测中的权重。显而易见，对两个项目共同评分的用户数量不同导致 dev_{ij} 的可靠度不同。假设有 500 个用户对项目 i 和项目 j 都产生过评分，而只有 50 个用户对项目 i 和项目 k 都产生过评分，那么显然 dev_{ij} 比 dev_{ik} 更可靠，dev_{ij} 应该以更高的权值参与到评分预测中。故文献[4]提出了加权 Slope One 算法(Weighted Slope One Algorithm, WSO)如公式(3)所示。

$$pre_{ui}=\frac{\sum_{j\in I(u)}(r_{uj}+dev_{ij})\cdot|U_{ij}|}{\sum_{j\in I(u)}|U_{ij}|} \tag{3}$$

Slope One 算法在计算两个项目的平均评分偏差进而进行评分预测时没有对不同用户和项目加以区分，导致推荐结果不够准确。盈艳等[5]基于 Slope One 算法预测评分矩阵中必要的未评分项，利用基于用户的协同过滤推荐算法实现推荐；Wang 等[6]先将用户聚类处理，然后对聚类后的用户群使用 Slope One 算法进行评分预测，提高了推荐质量；张玉连等[7]使用 Jaccard 系数找出具有影响力的用户和使用 Pearson 方法找出当前项目的相似项目对 Slope One 算法进行加权；刘林静等[8]通过评定用户活跃度的方式筛选出活跃用户，将活跃用户度作为评分预测的权重进而改善推荐准确度。然而上述方法均没有对项目和用户进行综合考虑，且都尚未对参与到评分预测中的用户已有评分进行有效区分。故本文针对这些问题，提出基于改进的项目相似性度量、改进的项目属性相似性度量和用户评分概率函数的多权值的 Slope One 协同过滤算法，对项目和用户进行了综合考虑，对不同的项目和用户进行有效的区分，改善了推荐系统对评分数据稀疏性的适应性，提高了推荐的准确度。

3 基于多权值的 Slope One 协同过滤推荐算法

3.1 基于改进的项目相似度度量的 Slope One 算法

Weighted Slope One 算法仅是简单地选取同时对两个项目都产生过评分的用户数量作为权重，在进行评分预测时并没有考虑到项目之间的差异，导致相似性较低的项目对评分预测产生了干扰，而相似性较高的项目没能有效参与到评分预测中。协同过滤旨在排除海量数据中不利信息对推荐系统的干扰，而对有利信息加以充分利用。故在使用 Slope One 算法产生推荐时，项目相似性较高的项目应该以更高的权值参与到评分预测中，相似性较低的项目应该以一个较低的权值参与到评分预测中，将项目相似性作为一个合理的权值参与到评分预测能够有效提高推荐效果。传统的相似性度量方法有余弦相似性、Pearson 相关系数、改进的余弦相似性[9]，其中用 Pearson 相关系数进行项目评分相似性度量如公式(4)所示。

$$sim_p(i,j)=\frac{\sum_{u\in U_{ij}}(r_{ui}-\overline{r}_i)(r_{uj}-\overline{r}_j)}{\sqrt{\sum_{u\in U_{ij}}(r_{ui}-\overline{r}_i)^2}\sqrt{\sum_{u\in U_{ij}}(r_{uj}-\overline{r}_j)^2}} \tag{4}$$

其中，$\overline{r}_i$ 和 $\overline{r}_j$ 分别为项目 i 和项目 j 的评分均值。

但是，在大多数推荐系统中，由于用户评分的项目往往不超过总项目数的 1%[10]，评分数据的稀疏不能准确地获取用户偏好，影响了协同过滤推荐算法的性能，是协同过滤推荐面临的一大挑战[11]。在用户评分数据较多的情况下，用上述 Pearson 相关系数度量项目评分相似性一般都能取得不错的度量效果，但是在评分数据稀疏的情况下用 Pearson 相关系数则不能准确地计算出项目之间的相似性，因为 Pearson 相关系数度量项目评分相似性将对两个项目都产生了评分的用户作为计算的集合，假设这个集合很小而且对两个项目都产生了评分的用户对两个项目的评分十分相近，所得的项目相似性会偏高，但实际上这是不可靠

的。既然对两个项目都产生评分的用户数量较小，直接将由这些用户评分计算得到项目相似性作为评分预测的权重就会产生较大的误差，因为少量的用户评分具有较大的偶然性，不具备较强的说服力。

本文针对 Slope One 算法的适用场景，将共同评价了两个项目的用户数量和 Pearson 相关系数度量方法相融合，提出一种改进的项目评分相似性度量方法，计算方法如公式(5)所示。

$$sim_r(i,j) = sim_p(i,j) \cdot \frac{|U_{ij}|^2}{|U_i| \cdot |U_j|} \tag{5}$$

其中，U_{ij} 表示对项目 i 产生过评分的用户数，$|U_{ij}|$ 表示同时对项目 i 和项目 j 产生过评分的用户数。

将其作为 Slope One 算法的一个权值参与到评分预测，得到基于改进的项目相似度的 Slope One 算法(Improved Item Similarity based Slope One Algorithm, ISSO)的评分预测方法，如公式(6)所示。

$$pre_{ui} = \frac{\sum_{j \in I(u)} (r_{uj} + dev_{ij}) \cdot sim_r(i,j)}{\sum_{j \in I(u)} sim_r(i,j)} \tag{6}$$

公式(5)考虑了共同评分项目数对项目相似度的影响，在共同评分项过少且评分较为接近的情况下，可以避免直接使用 Pearson 相关系数度量导致相似度过高的问题。因此将公式(5)中改进的项目相似性作为一个权值引入到 Slope One 算法中可以缓解评分数据稀疏情况下直接使用 Pearson 相关系数度量不够准确的问题，能够将项目相似性以一个更合理的权值参与到评分预测，使评分预测更为准确，有效提高推荐的准确度。

3.2 基于改进的项目属性度量的 Slope One 算法

在推荐系统中，不管是电子商务系统或是社交网络所涉及到的项目往往都会有类别信息，一个项目往往同时属于几个类别，例如一部电影可以同属于爱情片和喜剧片。传统的基于项目评分的相似性度量忽略了项目之间往往存在的属性联系，很显然两个项目属性类别交集越多那么它们就越类似，将项目属性相似性作为一个权值引入到 Slope One 算法中可以弥补项目评分相似性度量不够准确时给推荐系统带来的问题，提高推荐质量。传统的计算项目属性相似度的方法为 Jaccard 系数，即将项目属性的交集与并集的比作为项目的属性相似性。但是在实际应用中项目的属性往往较少，项目属性的交集也就更小甚至是没有交集，这就导致了直接使用 Jaccard 系数进行项目属性相似性度量时出现较多的零值，造成参与到评分预测中的项目属性相似性为零，这显然过于绝对，不能因为两个项目之间没有共同的属性就视项目属性相似性为零，这样会对评分预测产生较大的偏差。针对在度量项目属性相似性时直接使用 Jaccard 系数出现的零相似性问题，本文引入修正的拉普拉斯平滑(Laplace Smoothing)，提出一种新的项目属性相似性度量方法如公式(7)所示。

$$sim_a(i,j) = \frac{|A_i \cap A_j| + \alpha}{|A_i \cup A_j| + N \cdot \alpha} \tag{7}$$

其中，$|A_i \cap A_j|$ 和 $|A_i \cup A_j|$ 分别表示两个项目属性的交集和并集，N 的取值为两个项目交集的最大值，α 为可调参数。

拉普拉斯平滑主要是为了解决概率统计中某些分量在样本中没有出现过而导致的零概率问题，常见的方式为分子直接加 1，分母加上样本容量的大小。在概率统计中，假设样本很大，对统计结果的计数加 1 造成的概率变化可以忽略不计，然而推荐系统中两个项目属性交集较小，直接加 1 显然会带来较大的误差。所以针对推荐系统的适用场景，在原有 Jaccard 系数计算公式中设置一个可调参数 α ($0 \leqslant \alpha \leqslant 1$)，通过选取合适的 α 值计算项目属性相似性。则相应的基于项目属性的 Slope One 算法(Item Attributes based Slope One Algorithm, IASO)的评分预测方法如公式(8)所示。

$$pre_{ui} = \frac{\sum_{j \in I(u)} (r_{uj} + dev_{ij}) \cdot sim_\alpha(i,j)}{\sum_{j \in I(u)} sim_\alpha(i,j)} \tag{8}$$

3.3 基于用户评分概率函数的 Slope One 算法

在公式(3)中，对于每个目标用户已有的评分是不加区分的，即公式中每个 r_{uj} 的权值都是 1，这就忽略了用户的评分习惯及尺度的问题，在一定程度上影响了推荐的质量。有些用户偏向于打高分，有些用户则偏向于打低分，偏向于打高分的用户评分中出现的较低分不应该代表用户在大多数情况下给出的评分，参与到评分预测中的权重应该适当减少，偏向于给低分的用户的情况也类似。对于多数情况，在用户所有历史评分中出现较多的评分应该以更大的权重参与到评分预测中，因为这部分评分代表了用户更倾向给出的评分，有较大的概率出现。所以对于目标用户的不同评分，在评分预测中应该有不同的权重。本文将用户 u

对项目评分值为 *rate* 的概率表示为 p_u^{rate}，则其计算方法如公式(9)所示，用户评分概率函数计算方法如公式(10)所示，则相应的基于用户评分概率函数的 Slope One 算法(User Rating Probability Function based Slope One Algorithm, UPSO)的评分预测方法如公式(11)所示。

$$p_u^{rate} = \frac{|r_u^{rate}|}{|r_u|} \tag{9}$$

其中，$|r_u^{rate}|$表示用户 u 评分中评分为 *rate* 的个数，$|r_u|$表示用户 u 对所有项目产生的评分个数。

$$f(p_u^{rate}) = (p_u^{rate})^{\beta} \tag{10}$$

$$pre_{ui} = \frac{\sum_{j \in I(u)} (r_{uj} + dev_{ij}) \cdot f(p_u^{r_{uj}})}{\sum_{j \in I(u)} f(p_u^{r_{uj}})} \tag{11}$$

3.4 基于多权值的 Slope One 算法

上述 ISSO 算法可以在一定程度上缓解评分数据稀疏对推荐系统的影响，引入项目属性可以有效平衡项目相似性度量不够准确情况下带来的不利影响，最后引入用户评分概率函数从用户评分尺度和个人偏好的角度调整每个用户已有评分的权重，从用户和项目两方面综合考虑，可使评分预测更加趋于合理，提高推荐系统的稳定性和推荐质量。故将上述公式(5)、公式(7)、公式(10)和公式(2)组合可以得到基于多权值的 Slope One 算法(Multi-Weights based Slope One Algorithm, MWSO)，其评分预测方法如公式(12)所示。

$$pre_{ui} = \frac{\sum_{j \in I(u)} (r_{uj} + dev_{ij}) \cdot sim_r(i,j) \cdot sim_\alpha(i,j) \cdot f(p_u^{r_{uj}})}{\sum_{j \in I(u)} sim_r(i,j) \cdot sim_\alpha(i,j) \cdot f(p_u^{r_{uj}})} \tag{12}$$

4 实验及分析

4.1 数据集

本文采用的数据集为 MovieLens 站点提供的来自不同时间段的用户对电影评分的 ml-100k 数据集，该数据集记录了有过浏览行为的 943 个用户对 1 682 部电影的历史评分信息，并且提供了电影名称、电影类型、上映时间等电影属性信息，其评分数据稀疏度为 93.7%。本文实验选取的项目属性为电影类型，电影类型有科幻、冒险、动作、喜剧等 18 种类型，一部电影可以有多个类型。本文将该站点提供的数据集按 4∶1 的比例分成训练集和测试集，通过训练集的评分数据预测测试集中已有的评分。

4.2 评价指标

在评判推荐算法的准确度方面，通常把平均绝对偏差(Mean Absolute Error, MAE)[12]作为其指标。MAE 通过计算从训练集预测得到的用户评分与测试集中用户的真实评分的平均绝对差值进行推荐算法的准确性衡量。假设预测得到的用户评分为$\{s_1, s_2, \cdots, s_i\}$，对应的用户真实评分为$\{t_1, t_2, \cdots, t_i\}$，则 MAE 的计算方法如公式(13)所示。

$$MAE = \frac{\sum_{i=1}^{N} |s_i - t_i|}{N} \tag{13}$$

4.3 结果及分析

(1) ISSO 算法与 SO 算法对比实验

为了验证 ISSO 算法的推荐准确度，实验将评分数据按用户个数为 100、200、…、900 分成 9 组，分别使用ISSO算法和SO算法在不同用户数下进行评分预测，各自计算其 MAE 值，所得的实验结果如图 1 所示。

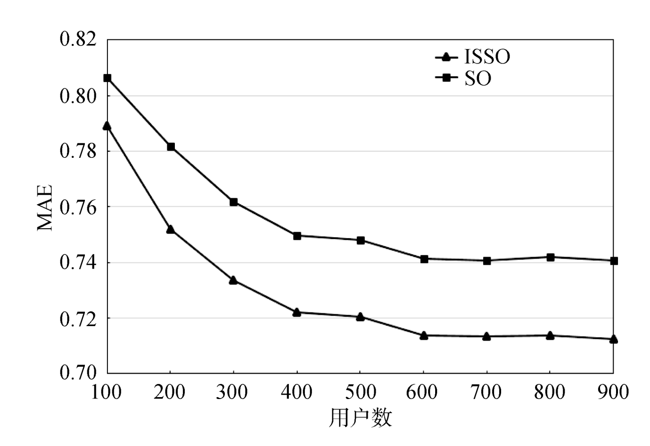


图 1 ISSO 算法与 SO 算法对比实验结果

通过对比实验的结果可以看出在不同用户数下 ISSO 算法的 MAE 值均明显低于 SO 算法，说明 ISSO 算法的预测准确度要高于 SO 算法，能够获得好的推荐效果。

(2) IASO 算法与 SO 算法对比实验

本文提出的 IASO 算法中α $(0 \leqslant \alpha \leqslant 1)$为可调节参数，针对推荐系统的适用场景，每次实验将参数α递增 0.1 进行评分预测并计算 MAE 值进行比较。作为对比的 SO 算法不含参数α，其 MAE 无变化趋势，所得的实验结果如图 2 所示。

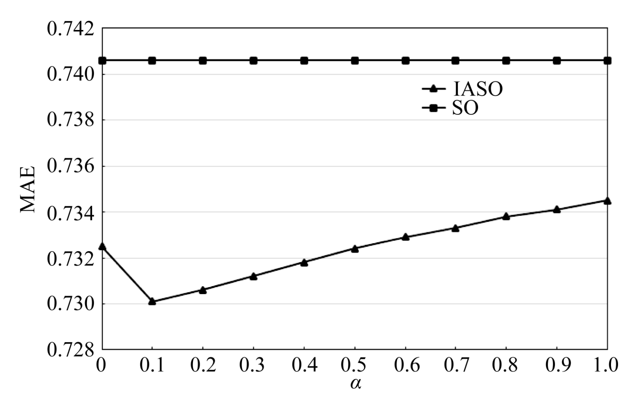


图 2　IASO 算法与 SO 算法对比实验结果

分析实验结果可知，当 α 为 0 时，项目属性相似度的度量方式变为 Jaccard 系数度量，计算所得 MAE 值较大，所以 α 为 0 不是最好的取值。当 α 的取值从 0 到 0.1 时，MAE 值呈下降趋势，推荐准确度有所提升。当 α 从 0.1 逐渐增大，MAE 值呈上升趋势，推荐质量下降。当 α 为 1 时，项目属性相似度的度量方式变为引入没有修正的拉普拉斯平滑及加一平滑的项目相似性度量，其 MAE 值较高，推荐效果也不理想。可见 α 取值过小或者过大都不利于项目属性以合适的权重参与到评分预测中，当 α 值为 0.1 时计算所得的 MAE 值最小，推荐准确度最高。

(3) UPSO 算法与 SO 算法对比实验

与实验(2)类似，将公式(9)中参数 β 的取值范围设为[0.1, 0.9]，每次实验将参数 β 递增 0.1 进行评分预测并计算 MAE 值。同样将 SO 算法作为对比，所得的实验结果如图 3 所示。

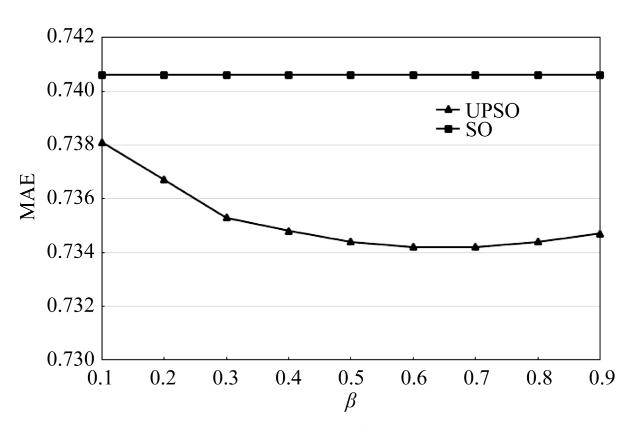


图 3　UPSO 算法与 SO 算法对比实验结果

从实验结果可知，当 β 取值为 0.6 左右时 UPSO 算法的 MAE 值最小，推荐准确度最高。结合用户评分概率函数以评分出现概率的形式区分用户的评分，使用户的评分以一个合理的权值参与到评分预测，使推荐更加准确。

(4) MWSO 算法与 SO 算法及其改进算法对比实验

与实验(1)类似，将评分数据按用户个数为 100、200、…、900 分成 9 组，分别设置参数 α 和参数 β 为 0.6 和 0.1，使用 MWSO 算法进行评分预测，并用 SO 算法、WSO 算法及文献[7]提出的融合用户相似度与项目相似度的加权 Slope One 算法中的 MSO2 算法进行对比，所得的实验结果如图 4 所示。

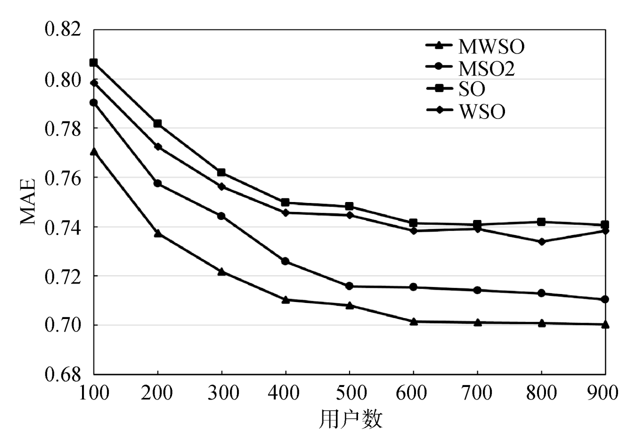


图 4　MWSO 算法与 SO 算法及其改进算法对比实验结果

从对比实验可以看出，在不同的用户数下 MWSO 算法相比 SO 算法、WSO 算法及 MSO2 算法其 MAE 值下降明显，即 MWSO 算法可以得到更好的推荐准确度。可见结合了项目相似性、项目属性和用户评分概率的 MWSO 算法可以综合用户和项目因素优势互补，且当其中一方度量不够准确时另一方可以对参与到评分预测中的权重进行平衡，使推荐系统在多种不确定因素面前保持性能的相对稳定，保证推荐质量。

5　结　语

本文研究的 Slope One 算法相比传统协同过滤算法具有运用简单快捷、对新用户和新项目适应性强和运行效率高等优点，但在计算过程中未考虑项目相似性、项目属性及未区分用户已有评分，造成推荐准确度缺失。针对这些问题，本文一共提出了 4 种方法，即基于改进的项目相似度度量的 ISSO 算法、基于改进的项目属性相似度度量的 IASO 算法、基于用户评分

概率函数的 UPSO 算法和将上述三种算法进行组合的基于多权值的 MWSO 算法。实验结果表明，上述从用户或者项目角度改进后的算法相较于原始 SO 算法均可得到更高的推荐准确度，综合了用户和项目因素的基于多权值的 MWSO 算法的实验效果最为突出，相比原始 SO 算法 MAE 下降了 5.4%，有效改善了推荐系统对评分数据稀疏性的适应性，提高了推荐准确度。然而本文提出的算法中都只关注了用户对项目产生的评分，没有关注用户对项目给出的评论，造成一些关键性信息缺失。因此如何将用户的评分和评论更好地结合起来对用户产生推荐将是下一步的重要研究内容。

参考文献：

[1] 王鹏，王晶晶，俞能海．基于核方法的 User-Based 协同过滤推荐算法[J]．计算机研究与发展，2013，50(7)：1444-1451. (Wang Peng, Wang Jingjing, Yu Nenghai. A Kernel and User-Based Collaborative Filtering Recommendation Algorithm [J]. Journal of Computer Research and Development, 2013, 50 (7): 1444-1451.)

[2] 居斌，钱沄涛，叶敏超.基于结构投影非负矩阵分解的协同过滤算法[J]．浙江大学学报：工学版，2015，49(7)：1319-1325. (Ju Bin, Qian Yuntao, Ye Minchao.Collaborative Filtering Algorithm Based on Structured Projective Nonnegative Matrix Factorization[J]. Journal of Zhejiang University: Engineering Science, 2015, 49 (7): 1319-1325.)

[3] 孙光明，王硕，邹静昭．多因素复合度量的协同过滤推荐算法[J]．计算机应用研究，2015，32(10)：2896-2900. (Sun Guangming, Wang Shuo, Zou Jingzhao. Collaborative Filtering Recommendation Algorithm Measured by Compound Multiple Fators[J]. Application Research of Computers, 2015, 32(10): 2896-2900.)

[4] Lemire D, Maclachlan A. Slope One Predictors for Online Rating-Based Collaborative Filtering [C]//Proceedings of the 2007 SIAM International Conference on Data Mining, Newport Beach, California, USA. 2007.

[5] 盈艳，曹妍，牟向伟．基于项目评分预测的混合式协同过滤推荐[J]．现代图书情报技术，2015(6)：27-32. (Ying Yan, Cao Yan, Mou Xiangwei. A Hybrid Collaborative Filtering Recommender Based on Item Rating Prediction [J]. New Technology of Library and Information Service, 2015(6): 27-32.)

[6] Wang J, Lin K, Li J. A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Slope One Scheme [C]// Proceeding of the 8th International Conference on Computer Science Education. 2013.

[7] 张玉连，郇思思，梁顺攀．融合用户相似度与项目相似度的加权 Slope One 算法[J]．小型微型计算机系统，2016，37(6)：1174-1178. (Zhang Yulian, Huan Sisi, Liang Shunpan. Integrating User Similarity and Item Similarity into Weighted Slope One Algorithm [J]. Journal of Chinese Computer Systems, 2016, 37 (6): 1174-1178.)

[8] 刘林静，楼文高，冯国珍．基于用户相似性的加权 Slope One 算法[J].计算机应用研究，2016，33(9)：2708-2711. (Liu Linjing, Lou Wengao, Feng Guozhen. New Weighted Slope One Algorithm Based on User Similarity[J]. Application Research of Computers, 2016, 33(9): 2708-2711.)

[9] Finkenzeller K. RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification [M]. Hoboken: John Wiley & Sons, 2003.

[10] 邹永贵，望靖，刘兆宏，等．基于项目之间相似性的兴趣点推荐方法[J]．计算机应用研究，2012，29(1)：116-118. (Zou Yonggui, Wang Jing, Liu Zhaohong, et al. Point of Interest Recommendation Method Based on Similarity Between Items[J]. Application Research of Computers, 2012, 29(1): 116-118.)

[11] 胡勋，孟祥武，张玉洁，等．一种融合项目特征和移动用户信任关系的推荐算法[J]．软件学报，2014，25(8)：1817-1830. (Hu Xun, Meng Xiangwu, Zhang Yujie, et al. Recommendation Algorithm Combing Item Features and Trust Relationship of Mobile Users[J]. Journal of Software, 2014, 25(8): 1817-1830.)

[12] Herloker J L, Konstan J A, Terveen L G, et al. Evaluating Collaborative Filtering Recommender Systems[J]. ACM Transactions on Information System (TOIS), 2004, 22(1): 5-53.

作者贡献声明：

覃幸新：提出研究思路，采集分析数据；

覃幸新，王荣波：进行实验，起草及修改论文；

王荣波，黄孝喜，谌志群：修改论文。

利益冲突声明：

所有作者声明不存在利益冲突关系。

支撑数据：

支撑数据由作者自存储，E-mail: 15957191408@163.com。

[1] 覃幸新．ml-100k．u.data．实验数据集．http://pan.baidu.com/s/1nu8gB49.

收稿日期：2017-04-26
收修改稿日期：2017-05-21

Slope One Collaborative Filtering Algorithm Based on Multi-Weights

Qin Xingxin Wang Rongbo Huang Xiaoxi Chen Zhiqun
(Institute of Cognitive and Intelligent Computing, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: [**Objective**] This paper aims to increase the recommendation accuracy with the help of modified Slope One algorithm. [**Methods**] We proposed a Slope One Collaboration Filtering Algorithm based on multi-weights, which improved the items' similarity measure, attributes similarity measure and users' rating probability function. Then, we combined the items' similarity measure with the number of users and Pearson correlation coefficient, the items' attributes similarity measure with modified Laplacian smoothing and Jaccard coefficient. We also identified users' ratings with a new probability function. [**Results**] The proposed method reduced the MAE by 5.4%, which increased the recommendation accuracy. [**Limitations**] The new method did not examine the users' comments, which might pose some negative effects to the recommendation accuracy. [**Conclusions**] The proposed algorithm could effectively improve the service of recommendation systems.

Keywords: Collaborative Filtering Slope One Multi-Weights Item Similarity Item Attributes

---

Delta Think 和 LibLynx 合作简化开放获取数据分析工具的访问

Delta Think 和 LibLynx 于近日宣布合作，进一步增强 Delta Think 开创性的开放获取数据分析工具(Open Access Data Analytics Tool, OADAT)的用户体验。

OADAT 将 Delta Think 的知识和专长与强大的数据分析和可视化功能相结合，为 OA 出版市场提供前所未有的见解。利用 Delta Think 汇总、组织、标记参照引用，并不断更新的对开放获取影响的评述，以及可执行细分切片和切割信息的交互式工具，组织可以深入了解与之直接相关的信息。

LibLynx 的身份和访问管理解决方案使得 Delta Think 的 OADAT 能够快速打包成为一个企业订阅服务。LibLynx 的基于 Web 的管理工具使得通过一系列流行的机构身份验证方法设置和配置基于订阅的访问变得快速简单，简化了用户的访问。

Delta Think 的首席执行官 Ann Michael 表示："将 LibLynx 访问管理部署到 Delta Think OA 数据分析工具，使发布商和机构的无缝访问成为现实。这种合作关系使我们的用户能够轻松地在其组织内分享洞察和数据分析，鼓励数据驱动的战略发展和对标比较文化。"

(编译自: http://www.liblynx.com/liblynx-delta-think-partner-to-simplify-institutional-access-to-oadat/)

(本刊讯)